



The history and fate of the gold standard

Citation

Jones, David S, and Scott H Podolsky. 2015. The History and Fate of the Gold Standard. The Lancet 385, no. 9977: 1502–1503. doi:10.1016/s0140-6736(15)60742-5.

Published Version

doi:10.1016/S0140-6736(15)60742-5

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:30367400>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

The History and Fate of the “Gold Standard”

David S. Jones and Scott H. Podolsky

For the past half century, physicians and clinical researchers have remained confident that randomized controlled trials (RCTs) provide the most rigorous test of preventive, diagnostic, and therapeutic interventions. They are ubiquitously referred to as the “gold standard” of empiric biomedical investigation, to the point where this status is often presented as a self-evident starting point in diagnostic or therapeutic evaluation. However, this status has long been contested, ever more so now with the emergence of “big data,” randomized registry trials, and other modes of knowledge production in medicine. In an era of increasing methodological self-reflection, it is useful to step back and examine how and when RCTs became the gold standard and what our aspirations for gold standards reveal about our deeper medical identity.

RCTs had a complicated pre-history, entailing attempts to ensure equivalent active and control groups, the occasional blinding of researchers or subjects, and the development of statistical methods of comparison. They took on their recognizably modern form with the British Medical Research Council’s landmark 1948 trial of streptomycin for pulmonary tuberculosis. As statisticians and clinical pharmacologists attempted to make sense of the post-World War II pharmaceutical revolution, the power of RCTs seemed critical. When the 1962 amendments to the Food and Drug Administration mandated proof of efficacy through “well-controlled” studies – namely, RCTs – prior to new drug approval, the United States government set the stage for the avalanche of pharmaceutical trials that followed. In Britain throughout the 1960s and 1970s, Archibald Cochrane advocated for the utility of RCT’s to sort therapeutic wheat from chaff. His work set the stage for such worldwide champions of rational therapeutic assessment as Thomas Chalmers, Iain Chalmers, David Sackett, and their colleagues.

But when did RCT’s become the *gold standard*? The first instance we have found of the phrase “gold standard” to refer to RCTs came in the pages of the *New England Journal of Medicine* in December 1982, in an article written by Alvan Feinstein and Ralph Horwitz. This date surprised us (and many of our colleagues) as a very late date for the first usage. Despite extensive searching, we have found no earlier occurrence of “gold standard” in reference to RCTs. We are eager to be proven wrong, but until all textbooks, conference proceedings, journals, and archival collections have been digitized and made full-text searchable, the gold standard of historical research itself remains elusive. Of interest, Feinstein and Horwitz described RCTs not as a gold standard that all research must strive to attain, but as an elusive ideal. Their special article was actually a brief in support of the rigorous conduct of other clinical epidemiological research designs. As they remarked, “epidemiological research has become increasingly important because it offers a substitute for the unattainable scientific gold standard of a randomized experimental trial.”

Given these surprises, it is worth looking more closely at how and when the concept entered the medical literature. The phrase “gold standard” has a long pre-history. It first appeared in *Lancet* in 1870 in a discussion of international coinage and efforts to restore the value of the guinea. Over the next sixty years it occurred repeatedly, in discussions of the actual gold standard: the technique of international finance that links the value of a nation’s currency to a set amount of gold, facilitating exchange between different

currencies. In the 1930s “gold standard” gained a new usage, in discussions of pharmacological use of gold, whether for tuberculosis (unsuccessfully, in 1934) or rheumatoid arthritis (successfully, in 1937). It first appeared in *NEJM* in 1933, in a humorous riff by Harvey Cushing about the state of surgery and dentistry in the United States. The next five references, through 1959, all referred to the financial gold standard.

However, the financial gold standard was rarely seen as a “gold standard.” Instead it proved controversial for much of its history. Isaac Newton put Britain on a gold standard in 1717, an arrangement formalized by the Royal Mint in 1816. Many other countries followed suit. The system broke down during the economic turmoil of the early twentieth century. Inflation during World War I forced England and other countries off the gold standard. A variant was restored in 1925, but that too had to be abandoned in 1931 during the Great Depression. After preliminary moves by the United States in the 1930s, Richard Nixon finally took the U.S. off its gold standard in 1971. In 1976 the U.S. government revised its definition of the dollar to remove all references to gold.

It was, ironically, in this setting – of the final abandonment of the financial gold standard in the U.S. – that the phrase began to appear in *Lancet* and *NEJM* as something valuable, not merely as a standard of exchange but as the definitive exemplar of quality and reliability. A 1975 *Lancet* review of new diagnostic criteria described how they set the “gold standard,” providing a new “esperanto of liver disease.” In *NEJM* in 1979 Victor McKusick described the presentations given by residents at Grand Rounds at John Hopkins as the “gold standard” for medical communication. Book reviewers described new textbooks as the “gold standard” for their fields. By the early 1980s, clinical researchers described specific procedures as the diagnostic or therapeutic gold standard (e.g., adrenal vein catheterization in *Lancet* in 1980, cardiac catheterization in *NEJM* in 1981, or hemodialysis in *Lancet* in 1982). After the first occurrence in 1982 of gold standard in association with RCTs, the phrase became commonplace, appearing less often within quotation marks, and paralleling the rise of the term “evidence-based medicine,” which first appeared in the medical literature in 1991.

What can we make of the irony of this usage entering medicine in the years after it was abandoned as a tool of international finance in the 1970s? It appeared in diverse therapeutic and diagnostic contexts, reflecting the broad aspirations in medicine for evidentiary solid ground and standardization throughout this era. But many of its uses in relation to RCTs were critiques, reflecting a legacy of the controversies that had long ensnared those who would claim the epistemic hegemony of RCTs. The debates about RCTs, and about the notion of medical “gold standard” more generally, often took on religious overtones. Angry about cardiologists’ demands that coronary artery bypass grafting be subjected to RCTs, Lawrence Bonchek encouraged surgeons in 1979 to “resist the almost religious fervor of those who would sanctify randomized studies as the only means of learning the truth.” Writing in 1992, P. Finbarr Duggan complained that the phrase “gold standard” itself “smacks of dogma” and should be abandoned.

The religious language here may not be coincidental. Arthur Kleinman and others have argued that the emergence of biomedicine within the monotheistic traditions of Europe and the Middle East imbued medicine with a commitment to universal truths, unitary paradigms, and a “single-minded approach to illness and care.” The idea of a gold standard, that there is one best way to do something, whether conduct clinical research, diagnose a disease, or treat a patient, emerges from this underlying commitment. While the

desire to base clinical decisions on the best possible evidence reflects a genuine effort to improve the quality of medical care, commitment to a gold standard does more than that. Allegiance to a single approach provides a focus around which communities can organize and rally. But critics have pointed to the dangers of such medical monotheism. As pioneering cardiac surgeon René Favaloro wrote in 1998, reflecting on three decades of debate about bypass grafting, “Randomized trials have developed such high scientific stature and acceptance that they are accorded an almost religious sanctification ... If relied on exclusively they may be dangerous.” Quoting Feinstein, Favaloro argued that medical decisions often had to be made without guidance from clinical trials: “To acknowledge this reality requires no loss of reverence, allegiance, or respect for the primacy of randomized trials as a ‘gold standard’ in scientific research.” While Favaloro saw this as a particular challenge for surgery, physicians in all specialties have oft-times resented the yoke of evidence-based medicine.

The past several years have seen increasing calls for an ecumenical approach to clinical research, with more flexible standards for what counts as acceptable study designs. Physicians have developed new methods to extract robust analyses from patient registries and from the ever-growing databases provided by electronic medical records. Will this erode the status of RCTs as a gold standard? The rise of personalized medicine, meanwhile, might make it more difficult to defend gold standards in diagnostic and therapeutic practice. Personalized medicine refocuses clinical attention away from the ‘typical’ patients analyzed by RCTs and onto the idiosyncrasies, genetic or otherwise, of individual patients. Has the phrase outlived its usefulness in medicine? It is too soon to tell. Yet even as some physicians turn away from their commitment to medical gold standards, some politicians, newly wary about global financial turbulence, talk of restoring the financial gold standard. Gold standards, whether actual or figurative, represent structures of exchange and aspirations toward stability, despite developments that threaten both.

Further readings:

- Bonchek, Lawrence I. “Are Randomized Trials Appropriate for Evaluating New Operations?” *NEJM* 301 (1979): 44-45.
- Chalmers, Iain. “The Development of Fair Tests of Treatment.” *Lancet* 383 (2014): 1713-1714.
- Daly, Jeanne. *Evidence-Based Medicine and the Search for a Science of Clinical Care*. Berkeley: University of California Press, 2005.
- Duggan, P. Finbarr. “Time to Abolish ‘Gold Standard.’” *BMJ* 304 (1992): 1568-1569.
- Favaloro, René G. “Critical Analysis of Coronary Artery Bypass Graft Surgery: A 30-Year Journey.” *JACC* 31:4 Suppl B (1998): 1B-63B.
- Feinstein, Alvan R., and Ralph I. Horwitz. “Double Standards, Scientific Methods, and Epidemiologic Research.” *NEJM* 307 (1982): 1611-1617.
- Kleinman, Arthur. “What Is Specific to Biomedicine?” In *Writing at the Margin: Discourse Between Anthropology and Medicine*. Berkeley: University of California Press, 1995, pp. 21-40.
- Timmermans, Stefan, and Marc Berg. *The Gold Standard: The Challenge of Evidence-Based Medicine*. Philadelphia, PA: Temple University Press, 2003.